

Details Regarding Experimental assessment of SRD's visual function

Supplementary material to accompany:

Ostrovsky, Y., Andalman, A. and Sinha, P. (2006). Vision following extended congenital blindness. *Psychological Science*, Vol. **17**, No. 12, pp 1009-1014.

Yuri Ostrovsky, Aaron Andalman, Pawan Sinha
Department of Brain and Cognitive Sciences
Massachusetts Institute of Technology
77 Massachusetts Avenue
Cambridge, MA 02139
USA

yostr@mit.edu, andalman@mit.edu, psinha@mit.edu

We decided to focus on two important domains of visual skills: basic form perception and face perception. Since these domains span a range of task complexity, it was our hope that they would be effective for assessing SRD's visual skills. The specific tests we conducted were organized as follows:

Basic form perception

Shape matching across space: visual search

Shape matching across time: visual memory

Shape matching across transformations

Depth perception from monocular cues

Object segmentation

Visual illusions

Face perception

Face/Non-face discrimination

Face localization in natural scenes

Face identification

Gaze estimation

Gender classification

In what follows, we describe each of these tests and SRD's responses.

1. Basic Form Perception

Skills such as matching simple shapes, searching for them amidst distractors and compensating for appearance changes caused by image transformations, are the building blocks of complex scene-analysis and object-recognition abilities. Given their fundamental significance for spatial perception, we devoted the first half of our test battery to these basic form perception tasks.

1.1. Shape matching across space: Visual Search

We tested SRD's ability to search for designated shapes within a matrix of distractors. Three sample stimulus displays are shown in figure 1. The polygonal shapes spanned, on average, 4 degrees of visual angle horizontally and vertically. The polygons were drawn in bright saturated colors, and within each trial all the polygons were the same color. The target polygons included simple regular polygons, arbitrary convex polygons, and complex polygons with numerous concavities.

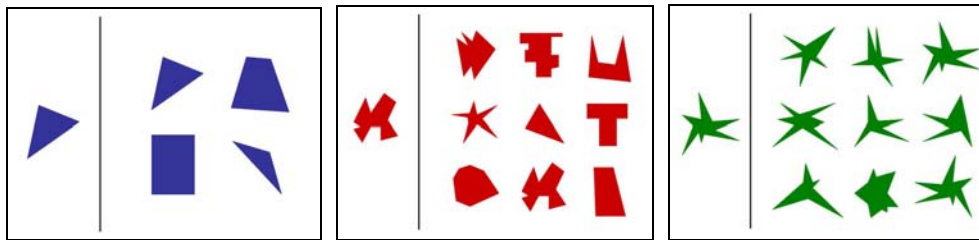


Figure 1. Three sample displays used to assess SRD's ability to match shapes and distinguish them from distractors. In each display, the designated target shape is shown on the left and the search matrix is one the right. SRD was asked to indicate whether the target appeared in the matrix (there were an equal number of target present/absent trials) and if so, to point to its location. SRD's performance was perfect on all trials.

Each trial proceeded as follows: The target polygon was shown alone on the left side of the screen and SRD was free to observe it for as long as she wanted. At her signal, a matrix of either 4 or 9 polygons appeared to the right of the target. SRD was asked whether the target was present within the matrix and, if so, to signal its location by touching it on the screen. SRD made no errors on all 11 trials ($p \ll 0.001$ compared to chance; χ^2 test), however we observed that her response times were slower than the control's. The status-matched control subject, who also made

no errors, was able to correctly respond within 2 seconds on each display, while SRD's response times ranged from 5 to 20 seconds. Acuity-matched controls also performed with no errors.

1.2. Shape matching across time: Visual Memory

As the temporal analogue of spatial shape matching, we tested SRD's ability to remember polygonal shapes over short periods of time. Two sets of polygons were displayed: a training set and a test set. SRD was allowed to study the training set without time restriction (typically ranging between 5 to 10 seconds per polygon), and was then asked to identify the training set members in the test set. The test sessions followed the training after a delay of 5 seconds. The polygons were similar to those used in the visual search task. They were uniformly colored and on average spanned 4 degrees of visual angle. Within each trial, all of the polygons were the same color. We ran three trials. The first two trials each had a training set of one polygon and a test set of five polygons, while the third had three training polygons and 15 test polygons. Figure 2 shows the overall design of this task.

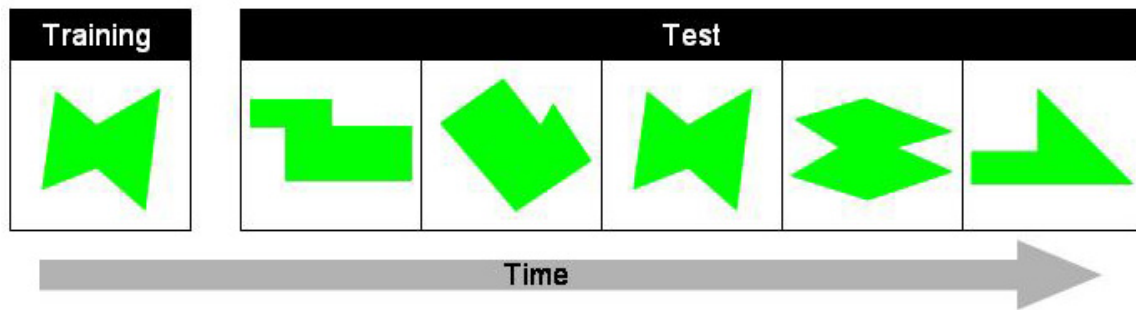


Figure 2. Task design for assessing SRD's ability to remember and match shapes presented across time. The training session was self-timed; SRD was allowed to look at the target pattern for as long as she wanted to. Upon presentation of the test polygons, SRD had to indicate whether they were the same as the previously seen target. SRD had a 100% hit rate and only two false positives.

SRD performed well above criterion in all three trials (hits: 8, misses: 0, correct rejections: 15, false alarms: 2; $d' > 2.4^*$). The status-matched control subject performed comparably, with just one false alarm ($d' > 2.8$), as did the three acuity-matched controls (perfect, 2 misses, and 1 false alarm, respectively).

* Due to the absence of misses, a direct computation of the d' score is not possible. Where the number of misses is 0, we have substituted 1. Therefore, our d' measurement is an *under-estimate* of the actual d' score.

1.3. Shape matching across transformations

In the tests described so far, the target did not undergo any change between the training and test phases. However, in the real world, an important challenge is to recognize objects even when they may have been subjected to transformations such as rotations and scale changes. We sought to determine whether SRD could compensate for transformations during shape matching.

Each display for this task consisted of a 2D or 3D geometric figure on the left side of the screen and two similar figures on the right side. SRD's task was to identify which figure on the right (top or bottom) matched the figure on the left, ignoring changes in size, color, orientation and (in the case of 3D shapes) lighting direction. Figure 3 shows a few stimulus displays.

SRD performed correctly on 24 out of 26 displays ($p \ll 0.001$; χ^2 test); the control subjects made no errors. (The difference in performance between SRD and the control subjects was not significant.) Notably, she answered incorrectly when discriminating between a pentagon and a hexagon whose orientations were manipulated (see upper-right panel in figure 3), perhaps due to the misleading similarity in orientation of the prominent peak of the pentagon and hexagon. She was not misled by similar manipulations of simpler shapes, such as a triangle versus a square.

SRD performed well on 3D shapes, not being misled by such confounds as in the lower-right panel of figure 3, where the illumination induced image-level similarities between the cube and the pentagonal prism are much greater than between the two cubes.

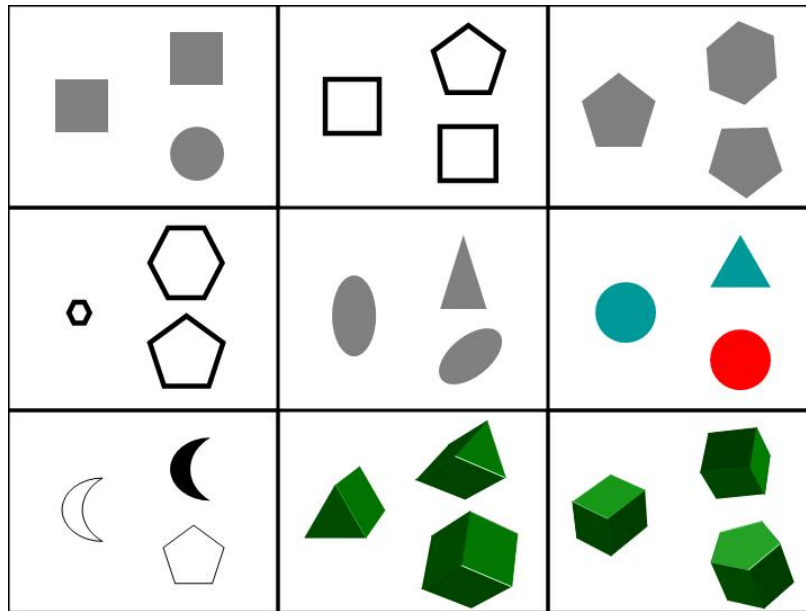


Figure 3. Sample displays from the shape matching across transformations experiment. For each display, SRD was asked to match the reference shape on the left with one of the two alternatives on the right, while ignoring transformations of scale, orientation, color or illumination. SRD matched 24 out of 26 trials correctly versus no errors by the control (see text for her mistakes).

1.4. Depth estimation from monocular cues

This series of 10 displays tested SRD’s ability to use image-based depth cues. The basic display consisted of two spheres in a virtual world with a textured ground plane (see figure 4). This was then modified through the addition of shadows, changes in image-level vertical location, changes in relative image-level size, and atmospheric haze. SRD was asked to indicate which sphere was “closer,” “bigger,” and “higher”. In all trials, SRD’s answers were consistent with a normal three-dimensional percept. That is, she appeared to be properly integrating image-based depth cues to form a 3D reconstruction of the image. She even properly used atmospheric lighting effects to determine relative nearness, despite a contradictory size-constancy cue. Responses of the control subject were identical to SRD’s.

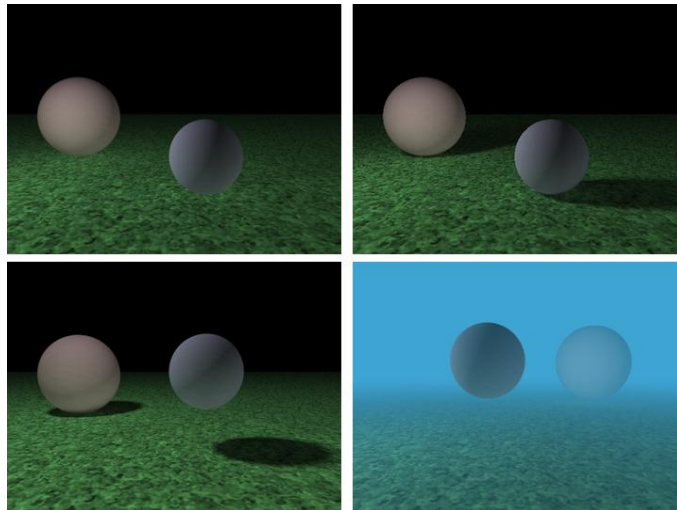


Figure 4. A few of the displays we used to assess SRD's ability to estimate depth based on monocular scene cues. SRD's assessment of all scenes was consistent with normal perception.

1.5. Object segmentation

So far, we have described tasks where the individual objects are entirely visible and figure-ground segregation is relatively straightforward. However, an important challenge in many settings is object segmentation – parsing a given image into coherent segments corresponding to distinct objects, even when they partially overlap and occlude each other. We tested SRD's object segmentation ability using two tasks. The first required SRD to point to and count the number of objects she perceived within each scene. The second required her to identify the partially overlapping objects present within each scene.

For the first task, we created 20 displays each consisting of two or three polygons with or without overlap. Each polygon spanned, on average, 7 degrees of visual angle. Two special trials were included in which photographs of overlapping pool-balls and wooden building blocks were used. Figure 5 shows some of the experimental displays.

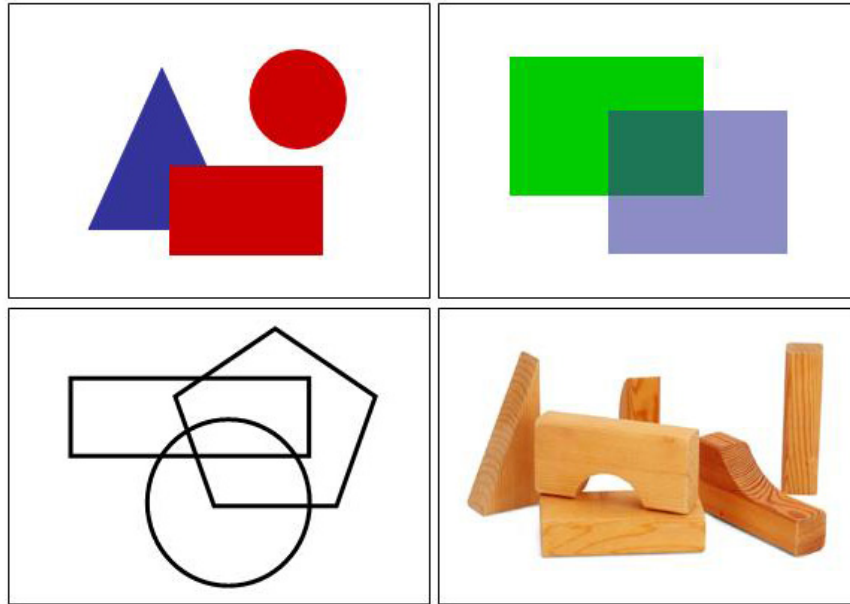


Figure 5. Four sample displays of the kind we used to assess SRD's image segmentation skills. For each display, SRD was asked to indicate the number of objects and point to their locations. SRD indicated the same objects as normal subjects on all trials.

SRD reported the same counts as the control in all 20 trials including those involving transparency, or unfilled line-drawings. Her success on the photographs of identically stained wooden building blocks was especially impressive given her acuity. This trial required the use of shading cues and small misalignments of bounding contours.

The second task required SRD to identify photographs of real-world objects such as trucks, cars, cups, saucers and spectacles. Each display comprised two objects with partial overlap. Figure 6 shows some of the experimental displays. SRD was able to correctly name all objects except a photograph of a fish, which she labeled an alligator. The control subject made no errors on either of the two tasks.

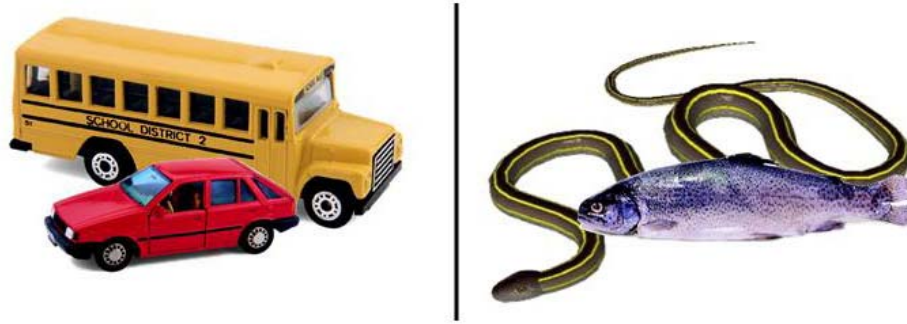


Figure 6. Two sample displays of the kind we used to assess SRD’s object naming skills in the presence of partial overlap. For each display, SRD was asked to name and point to all of the objects in the scene. SRD identified all objects that we tested, except the fish, which was labeled as an alligator.

1.6. Susceptibility to visual illusions

We concluded our testing of SRD’s basic form perception skills with a small battery of classic illusions. Our intent was to determine how congruent SRD’s responses were to those of normal observers. Significantly divergent responses would indicate differences in the underlying visual processing mechanisms. The specific tests we conducted were the Müller-Lyer illusion, inverted T illusion, 3-dot bisection hyperacuity and the simultaneous contrast illusion. Some of our displays are shown in figure 7. As described in the figure caption, SRD was susceptible to all of these illusions in a manner consistent with normal observers.

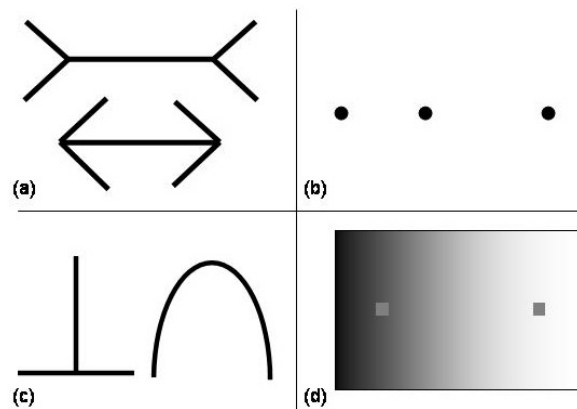


Figure 7. Sample displays from our tests of SRD’s susceptibility to some classic visual illusions. (a) Müller-Lyer illusion; SRD was asked to indicate which line appeared longer and was instructed to pay attention to the line only, and not the overall size of the object. (b) Bisection hyperacuity display. SRD was asked to say whether the middle dot was to the left or right of center. She performed perfectly down to a 5 pixel offset (corresponding to approximately 18

minutes of visual angle). (c) *Inverted T and arch illusion*. SRD was asked to indicate whether the T's vertical or horizontal line segment appeared longer, and whether the arch was greater in width or height. She consistently chose the T's vertical stem and indicated that the arch was taller than wide. (d) *Simultaneous-contrast illusion*. Even though the two small squares are of the same physical luminance, SRD, like normal observers, reported the square on the right as being darker.

Overall, SRD's performance on the basic object perception tasks included in our test battery is close to that of the control subject. While this does not mean that she does not have any residual impairments, this body of results does suggest that a significant level of proficiency on form-perception skills can be acquired even after several years of congenital visual deprivation. Our next set of experiments tested SRD's performance on more complex and ecologically relevant form-perception tasks, specifically the analysis of faces.

2. Face Processing

In the domain of face perception, we considered two broad classes of skills: (1) distinguishing between faces and non-faces (closely related to localizing faces in complex scenes), and (2) making use of intrinsic face information for assessing attributes such as identity, gender and gaze-direction. Our tests with SRD were designed to probe both of these kinds of skills.

2.1. Face/Non-face discrimination

SRD was presented with randomly interleaved grayscale face and non-face patterns and was asked to classify them as such. Images were shown one at a time and remained on the screen until SRD had responded verbally ('Yes' or 'no' to the question: Is this a face?). Our stimulus set comprised 30 patterns. Of these, 15 were faces of both genders in a frontal viewpoint, showing the face from the middle of the forehead to the chin. Each image subtended 7 degrees horizontally at a viewing distance of 30 cm. The set of non-face distracters comprised 15 patterns from natural images that had similar power-spectra as the face patterns and also false-positives of computational face-classification systems. The system we used is Rowley et al's (Rowley, H. A., Baluja, S. and Kanade, T. (1998). Neural-network based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **20**, 23-28.) scheme from CMU. Sample non-face images are shown in figure 8.



Figure 8. A few of the face (top two rows) and non-face (bottom two rows) patterns used in our experiments. The non-face patterns comprise false alarms of computational face-detection systems and images with similar spectra as face images. SRD had a 100% hit rate, with 1 false positive. The status-matched control too had a perfect hit rate with 1 false positive.

SRD achieved a perfect hit rate, correctly classifying all of the face patterns as faces. Her only mistake was one false-positive. The status-matched control subject too had a 100% hit rate and one false-positive.

2.2. Face localization in scenes

The face/non-face classification task described above is somewhat artificial in its design. In the real-world, we are typically faced with the task of locating faces in complex scenes. Success on this task depends not only on discriminating between faces and non-faces, but also on the use of additional information arising from the rest of the body and scene structure (Cox et al., *Science*, vol 303, 115-117, 2004). We assessed SRD's face localization skills on this more naturalistic task.

SRD was shown 10 full-color natural images containing people in diverse settings. She was asked to point to faces in these scenes. The localization responses were recorded by the experimenter and their veridicality assessed relative to pre-determined face-location data for each stimulus image. Figure 9 shows a few sample scenes of the kind we used in the experiments. Scenes subtended 30 x 22 degrees of visual angle at a viewing distance of 30cms. The scenes included faces in different viewpoints, ranging from profile to frontal. The widths of the faces ranged from 1 degree to 10 degrees. Across the entire stimulus set, we ensured that the faces were distributed evenly over the display area so as to avoid the development of positional biases in responses.

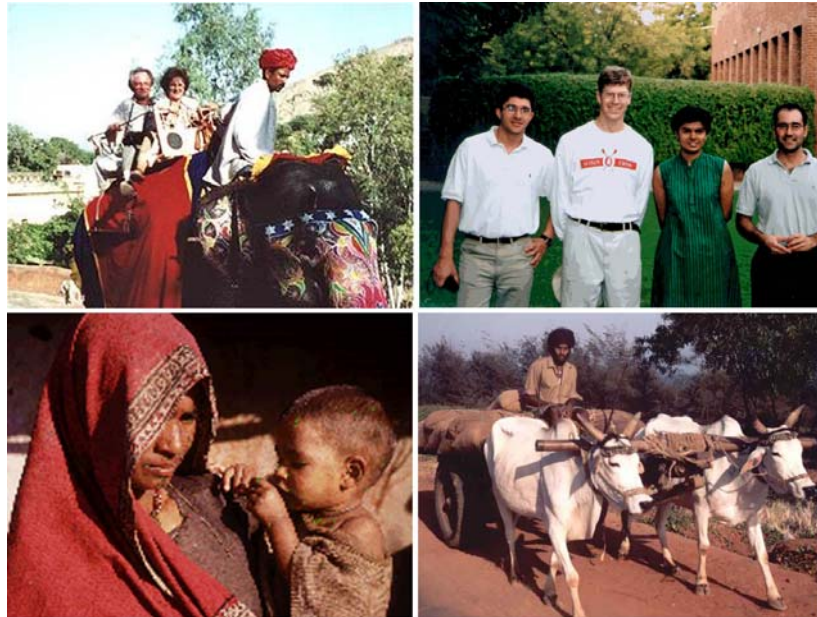


Figure 9. Sample scenes of the kind that we used in our study of face localization, wherein the subject was asked to point to all faces in the scene. SRD's performance was flawless on this task.

Both SRD and the control subject performed flawlessly on this task. SRD correctly pointed to all of the faces in the scenes and did not exhibit any false alarms.

Given SRD's impressive skills at generic face localization, we next assessed her ability to individuate people and determine their gender and gaze direction.

2.3. Face Identification

We probed SRD's face identification abilities using a set of nine face matching tasks. All nine tasks followed a delayed match to sample paradigm, and each comprised 3-5 trials. In the first seven tasks, SRD was shown a sample face to remember followed by two faces, one of the same person, the other of a different person. Both of the faces were transformed according to the task. SRD was then asked to select the face that was of the same person. The seven task transformations were: 1. vertical inversion, 2. occlusion of the upper-half, 3. occlusion of the lower-half, 4. 45 degree depth rotation, 5. Gaussian blur, 6. edge-detection (which transforms the images into black and white line drawings), 7. the luminance transform (changing the overall

luminance of the face), and 8. feature replacement (eyes and the mouth in one of the test faces were replaced by the eyes and the mouth of a randomly chosen face).

The faces used in the task were chosen from the face database compiled by researchers at the Max-Planck Institute for Biological Cybernetics in Tuebingen, Germany. The faces were displayed in color and subtended 10 degrees of visual angle horizontally. They were all clean shaven and had no artifacts that could aid identification (such as eye-glasses or moles). All faces were presented on a black background and were cropped at the hairline to prevent having the hair serve as an identity cue. Figure 10 shows some of our stimuli.

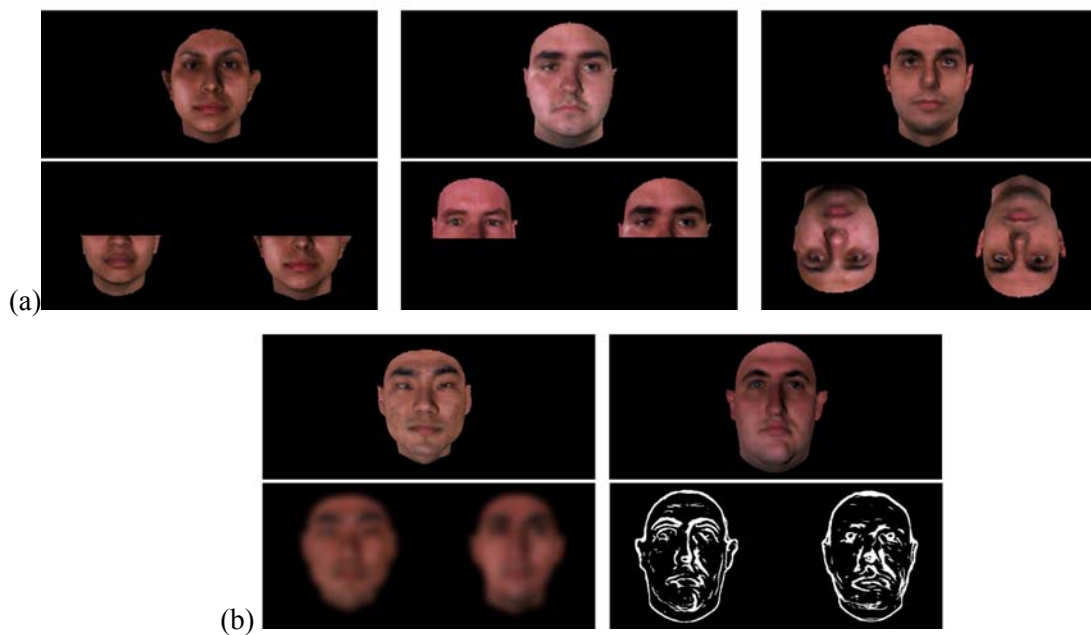


Figure 10. Sample stimuli used to assess SRD's face recognition performance. Shown here are five of the transformations we used – (a) occluding the upper half of the face, occluding the lower half and vertical inversion and (b) blur and edges-only. SRD was shown the target face in a self-timed fashion, and following a blank screen delay of approximately 2 seconds, the two alternatives were presented. SRD had to identify which of the two alternatives was the target face. The Gaussian blur filter we used had a radius of 14 pixels applied to face images with a pupil to pupil distance of 150 pixels. SRD's reduced acuity is akin to a Gaussian filter of approximately 4 pixel radius. The edge versions of our face stimuli involve non-linearities such as thresholding and manual noise removal, and, therefore, do not permit a simple filtering analog.

SRD's performance was as follows:

Vertical inversion:	5/5
Edges only:	3/3
Featural changes:	3/3
Upper-half occlusion:	4/5
Bottom-half occlusion:	4/5
Luminance changes:	3/5
45 degree depth rotation:	3/5
Gaussian blur:	3/5

The control subject, by comparison, performed at ceiling on all of these tests. Although we are clearly working with a limited stimulus set size, some trends are evident. SRD's performance, though not perfect, appears to be near normal for most transformations. Across the entire stimulus set, SRD's performance is significantly above chance ($p < 0.005$; χ^2 test). SRD's aggregate performance was 28/36 which was slightly worse than our acuity-matched controls at 34/36, 33/36, 31/36 and 33/36 ($p < 0.05$; χ^2 test). Our status-matched control probably performed better because her displays were not adjusted for acuity, suggesting that acuity may have been responsible for at least some of SRD's remaining deficit.

2.4. Gaze estimation

In our day to day interpersonal exchanges, we use faces to extract more information than just identity. An important example is direction of gaze. Estimating gaze direction is a pre-requisite for tasks that require shared attention. We assessed how well SRD was able to estimate facial gaze direction. Given her impressive performance on the previous face perception tasks, we expected her to be quite proficient at determining gaze direction. We showed SRD 20 images of individuals facing the camera, but shifting their eyes to look in one of three different directions: straight ahead, left, or right (see figure 11 for examples).

Contrary to our expectations, SRD performed very poorly on this task. In all cases, she reported gaze consistent with the orientation of the head and ignoring the orientation of the iris. Even slight depth rotations of the head coupled with exaggerated eye positioning in the opposite direction caused SRD to interpret gaze consistent with head pose and inconsistent with eye position. This was not due to an inability to resolve the irises in the images presented. At distances of about 1.5 m, SRD was able to correctly determine whether the pupil appeared to the

left or to the right of center. When told that this positioning corresponded to gaze direction, SRD was surprised and said that she was unaware of such a correspondence. Eye-gaze estimation, thus, is a skill that is highly compromised in SRD. For the control subject, this task was trivially easy and she performed without errors on all of the stimuli. SRD's reliance on head-orientation rather than the intra-ocular iris position as an indicator of gaze direction is perhaps a consequence of her compromised acuity. Poor acuity might render iris position irresolvable at distances typical of social interactions, while still permitting head orientation to serve as a useful cue.



Figure 11. Sample images of the kind we used to assess SRD's eye-gaze estimation skills. It was discovered that SRD used the orientation of the head, not of the eyes, as the cue to gaze orientation. She was neither perceptually nor cognitively aware of the relationship between eye orientation and gaze direction.

2.5. Gender classification

The final test of face perception we conducted asked SRD to specify the gender of 10 faces. The 10 images were evenly divided across the genders and showed faces cropped at the forehead and the chin. Thus, hair and clothing were not available as cues. None of the women in the set were wearing the traditional (for India) red dot on their forehead. SRD correctly classified all of the images. Her response times were comparable to those of the status-matched control subject.

Postscript:

Just as the paper was going to press, the researchers received some sad news. SRD met with an accident while taking her 9 year old daughter to the eye-clinic. She fell while alighting from a public-transport bus (the driver started the bus before she had fully stepped out) and was dragged under the wheels. She died from her injuries. The daughter, who was not injured in the accident, now lives by herself in a hostel for the blind. She has been disowned by the father due to financial constraints. The researchers will contribute funds to help her with the living/educational expenses. If you would like to help too, please contact any of the authors.